

Reading signatures of supermassive binary black holes in pulsar timing array observations

Received: 3 December 2024

Accepted: 15 October 2025

Published online: 03 November 2025



Boris Goncharov^{1,2}✉, Shubhit Sardana^{1,3}, Alberto Sesana^{4,5,6}, Sharon Mary Tomson^{1,2}, John Antoniadis^{7,8}, Aurelien Chalumeau⁹, David J. Champion⁸, Siyuan Chen^{10,11}, Evan F. Keane¹², Kuo Liu^{10,11}, Golam Shaifullah^{4,5,13}, Lorenzo Speri¹⁴ & Serena Valtolina^{1,2}

Constraining the origin of the nanohertz gravitational-wave background necessitates precise noise modelling to avoid parameter estimation biases. In this work, we find the inferred properties of the putative gravitational wave background in the second data release of the European Pulsar Timing Array to be in better agreement with theoretical expectations under the improved noise model. In particular, our improved noise models show consistency of the background's strain spectral index with the value of $-2/3$, favoring the population of supermassive black hole binaries as the origin of the background. Our results further suggest that the observed gravitational wave emission is the dominant source of the binary energy loss, with no evidence of environmental effects or eccentric orbits. At the reference gravitational wave frequency of yr^{-1} , we also find a lower power-law strain amplitude of the background than in previous data analyses. This mitigates some of the tensions of the strain amplitude with the expected number density and mass scale of binaries discussed in the literature. Our analysis demonstrates the importance of accurate modelling of radio pulsar pulse profile variations, hierarchical properties of noise across pulsars, as well as noise model averaging, when inferring properties of the gravitational wave background.

Since 2020, Pulsar Timing Arrays (PTAs) have reported growing evidence for the nanohertz-frequency gravitational wave background in their data. The first tentative evidence came from a temporally correlated stochastic process in pulsar timing data^{1–4}. The Fourier spectrum of delays and advances in pulsar pulse arrival times exhibited the expected spectral properties of the background. Most recently, PTAs –

with varying levels of statistical significance^{5–8} – showed that this stochastic process exhibits Hellings-Downs correlations⁹ consistent with the isotropic unpolarised stochastic gravitational wave background.

Properties of the gravitational wave background are assessed using the power law model of its characteristic strain spectrum:

¹Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Hannover, Germany. ²Leibniz Universität Hannover, Hannover, Germany. ³Department of Physics, IISER Bhopal, Bhopal, India. ⁴Dipartimento di Fisica “G. Occhialini”, Università degli Studi di Milano-Bicocca, Milano, Italy. ⁵INFN, Sezione di Milano-Bicocca, Milano, Italy. ⁶INAF - Osservatorio Astronomico di Brera, Milano, Italy. ⁷FORTH Institute of Astrophysics, N. Plastira 100, Heraklion, Greece. ⁸Max-Planck-Institut für Radioastronomie, Bonn, Germany. ⁹ASTRON, Netherlands Institute for Radio Astronomy, PD Dwingeloo, The Netherlands. ¹⁰Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China. ¹¹State Key Laboratory of Radio Astronomy and Technology, Chaoyang District, Beijing, P. R. China. ¹²School of Physics, Trinity College Dublin, College Green, Dublin 2, Ireland. ¹³INAF - Osservatorio Astronomico di Cagliari, Selargius, CA, Italy. ¹⁴Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Potsdam, Germany. ✉e-mail: boris.goncharov@me.com

$h_c(f) = A(f/\text{yr}^{-1})^{-\alpha}$. Here, A is the strain amplitude at the gravitational wave frequency f of $f_{\text{ref}} = \text{yr}^{-1}$, and $-\alpha$ is the power law spectral index. The corresponding spectral index of the power spectral density [s^3] of delays (advances) [s] induced by the background in the timing data is $-\gamma$. These stochastic timing delays resulting from the gravitational-wave redshift of pulsar spin frequency are referred to as temporal correlations.

Supermassive black hole binaries at subparsec separations are expected to be a dominant source of the stochastic gravitational wave background at nanohertz frequencies¹⁰. However, the expected amplitude of the background is lower than the observations suggest¹¹. This is visible in Figure 7 from ref. 12 and Fig. 5 from ref. 13, where the best-fit strain amplitude is at the edge of the values simulated from supermassive black hole binary population synthesis models. The inferred strain amplitude lies at the theoretical upper limit of the predicted astrophysical range^{10,14,15}. It might also be in tension with the observed black hole mass function^{16,17}, although see¹⁸ for a different view. Furthermore, the strain spectral index of the gravitational wave background is in about 2σ tension with the value corresponding to binary inspirals driven by gravitational wave emission alone ($\gamma = 13/3$, $\alpha = 2/3$)¹⁹. The tension is visible in Fig. 5 in ref. 13 and Fig. 11 in ref. 5, where the posterior on the spectral index is compared not with a single value, but with the distribution of values expected from realistic backgrounds made up of discrete point sources. Although previous PTA results were consistent with a very broad range of assumptions about binary black hole populations²⁰, they suggested deviations from purely gravitational wave-driven binary evolution. Furthermore, deviations of the strain spectral index from $-2/3$ can point to the early-Universe origin of the signal^{13,21}.

Modelling of pulsar-intrinsic noise is important because it can affect the conclusions about the properties of the gravitational wave background, such as A and γ ²². It is relevant for the European Pulsar Timing Array (EPTA), one of the world's leading PTAs, where the latest 10-year data set showed evidence for the gravitational wave background at 3.5σ ⁶. Despite dedicated noise modelling studies being performed for the EPTA²³, there are four indications that some noise is still mismodelled. First, it was pointed out in several studies^{2,4,24} that the standard PTA models of how noise parameters are distributed across pulsars are incorrect. These models manifest as prior probabilities in PTA data analyses. To be precise, the models are incorrect

because they are 'static', i.e., the shape of the distribution of noise parameters is not influenced by the data. Although imposing such priors is very unlikely to influence our conclusions about evidence for the gravitational background^{25–28}, it may introduce systematic errors in the inferred strain spectrum²⁹. Second, epoch-correlated temporally-uncorrelated (white) noise term³⁰ was neglected in the previous EPTA analysis. Third, van Haasteren³¹ pointed out that the procedure of removing certain noise terms in the search for gravitational waves based on the results of single-pulsar noise analysis, performed in the original EPTA analysis⁶, is prone to systematic errors. Fourth, the original EPTA analysis treats the transient noise effect in pulsar PSR J1713 + 0747 as from a sudden change in dispersion measure, although Goncharov et al.³² suggested that it is associated with the pulse shape change.

In this work, we address the above four limitations and revise the properties of the gravitational wave background inferred by the European Pulsar Timing Array (EPTA)⁶. As a result, we find a steeper characteristic strain spectrum of the background, which is in better agreement with the hypothesis that the background originates from a superposition of adiabatically inspiraling supermassive binary black holes in circular orbits. Based on the revised amplitude and spectral index of the background, we discuss implications for the evolution of supermassive black hole binaries. Finally, we describe the observational impact of our new noise model. In particular, we show that despite our new model results in less evidence for Hellings-Downs correlations, evidence remains strong. Our results highlight the importance of accurate noise models for correctly inferring background properties.

Results

Posterior distributions for A and γ are shown as contours in Fig. 1. Solid blue contours correspond to our improved model. For comparison, dashed red contours correspond to the noise model used in the original analysis of the EPTA data⁶. The results are shown for both the 10-year subset of the EPTA data, which showed evidence for the Hellings-Downs correlations, and the full 25-year EPTA data, where the evidence is not visible³³. The value of $\gamma = 13/3$ ($\alpha = 2/3$) is shown as a dashed straight line. Our improved model results in a lower median-*a posteriori* strain amplitude of the gravitational wave background, as well as in a steeper spectral index, as shown with solid blue contours in

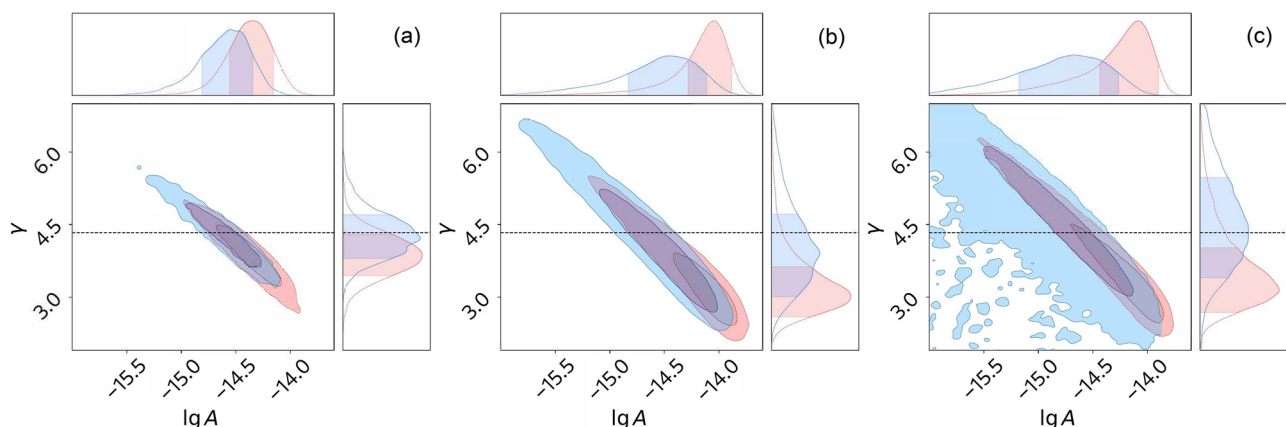


Fig. 1 | Posterior distributions for the power-law amplitude A and spectral index γ of the putative gravitational wave background in the European Pulsar Timing Array (EPTA) data. The results of a fit to Hellings-Downs correlations are shown in the left panel (a, full 25-year data, $(\lg A, \gamma) = (-14.53^{+0.15}_{-0.27}, 4.14^{+0.56}_{-0.26})$) and the middle panel (b, the 10-year subset of the data described in ref. 6, $(\lg A, \gamma) = (-14.37^{+0.17}_{-0.57}, 3.38^{+1.50}_{-0.25})$). The results of a fit of only temporal correlations to the 10-year data are shown in the right panel (c, $(\lg A, \gamma) = (-14.52^{+0.15}_{-0.80}, 4.20^{+1.32}_{-0.72})$). Dashed red contours correspond to the result using the standard pulsar

noise priors, and the blue contours correspond to our improved model. The horizontal dashed line corresponds to the background from supermassive binary black holes inspiralling entirely due to gravitational wave emission (subject to cosmic variance). Our improved model results in a lower median-*a posteriori* strain amplitude of the background and mitigates tensions with $\gamma = 13/3$. In marginalised distributions, shaded areas correspond to 1σ credible levels. In joint distributions, an inner dark area corresponds to the 1σ level, and the outer lighter area corresponds to the 2σ level.

Fig. 1. Maximum-*a-posteriori* ($\lg A, \gamma$) obtained with our improved model and the respective measurement uncertainties based on 1σ credible levels are reported in the caption of Fig. 1.

A detailed inspection of the contribution of four new components of our noise models, not shown in Fig. 1, reveals the following. Using noise model averaging instead of model selection, as recommended by ref. 31, has mostly affected the measurement uncertainty. Every other component of our improved noise model results in a consistently lower background strain amplitude and a steeper spectral index. Resolving noise prior misspecification using hierarchical inference²⁴, the first component of our improved noise model, has mostly affected the 25-year data (Fig. 1a) and the 10-year data when inter-pulsar correlations are not modelled (Fig. 1c). When modelling Hellings-Downs correlations in the 10-year data (Fig. 1b), other components of our revised noise model have made most of the impact. In particular, we found that the amplitude of the transient noise event in PSR J1713+0747 depends on the radio frequency consistently with ref. 32, not as assumed in the original EPTA analysis⁶. Modelling a transient noise event in PSR J1713 + 0747 correctly results in a significant shift in the posterior. Finally, epoch-correlated noise has not been considered in previous EPTA analyses because of a low number (1–4) of pulse arrival time measurements per observation epoch compared to other PTAs^{6,23}. Nevertheless, we find evidence for this noise in the data from certain backend-received systems. In Fig. 1b, the choice leads to the consistency of the fully marginalised posterior on γ with 13/3 at the 1σ level. Revising noise priors alone, without considering epoch-correlated noise, yields a weaker consistency of the joint posterior on ($\lg A, \gamma$) with $\gamma = 13/3$ at the 1σ level (not shown in Fig. 1b). Stronger impact of hierarchical inference in Fig. 1a, c compared to Fig. 1b illustrates how inter-pulsar correlations influence the posterior.

Implications for supermassive binary black holes

If the energy loss in binary inspirals is dominated by the adiabatic emission of gravitational waves and if binaries are circular, the characteristic strain spectrum of the gravitational wave background is ref. 19

$$h_c^2(f) = \frac{4G^{5/3}}{3\pi^{1/3}c^2} f^{-4/3} \int \frac{d^2N}{dVdz} \frac{\mathcal{M}^{5/3}}{(1+z)^{1/3}} dz, \quad (1)$$

where G is the Newton's constant, c is the speed of light, z is redshift, \mathcal{M} is the binary chirp mass, and $d^2N/(dVdz)$, a function of (\mathcal{M}, z) , is the number density of binaries per unit comoving volume per unit redshift. The integral does not depend on a gravitational wave frequency, thus $h_c \propto f^{2/3}$, as stated earlier. The background amplitude A depends on the mass spectrum and the abundance of supermassive binary black holes in the universe. The derived value of $\alpha = 2/3$ ($\gamma = 13/3$) is confirmed by population synthesis simulations, e.g., Figure 7 in ref. 12, where the theoretical uncertainty is only $\delta\gamma$ of approximately 0.1 at 1σ due to cosmic variance^{34,35}.

Tensions of γ of 3 inferred from the previous EPTA analysis with $13/3 = 4.3(3)$ reported during the announcement of evidence for the gravitational background^{5–8} has led to discussions on whether the signal is influenced by certain effects of binary evolution that make the strain spectrum to appear flatter. Mechanisms of flattening $h_c(f)$ typically involve the introduction of a more rapid physical mechanism of a reduction in binary separation compared to a gravitational wave emission at < 0.1 parsec separations. Such a mechanism could be an environmental effect³⁶, such as stellar scattering^{37,38}, the torques of a circumbinary gas disc³⁹. Furthermore, it could be due to the abundance of binaries in eccentric orbits, which leads to a more prominent gravitational wave emission⁴⁰. Although eccentricity also results in a steeper $h_c(f > 10^{-8}\text{Hz})$ ⁴¹, but PTA sensitivity declines towards high frequencies. In contrast, the results of our improved analysis are

consistent with binary evolution driven only by the emission of gravitational waves.

Our improved model also impacts the measurement of the strain amplitude, which is directly related to the number density of supermassive black hole binaries. The new results suggest that the supermassive black hole binaries are not as (over-)abundant as the earlier measurements implied. This is shown in Fig. 2, the bottom panel of which is based on Figure A1 in ref. 12. Grey horizontal bands correspond to theoretical uncertainties on the strain amplitude at the 16th–84th percentile level in 26 studies^{10,15,42–65}. For ref. 50 the model is “HS-nod”, and for ref. 51 the model is “HS_nod_SN_high_accr”. Our improved model reduces tensions of the gravitational wave background strain amplitude with theoretical and observationally-based predictions for supermassive black hole binaries. Furthermore, the revised background strain amplitude corresponds to longer delay times between galaxy mergers and supermassive binary black hole mergers⁵¹. The caveat is that the reported amplitude is referenced to $f = \text{yr}^{-1}$, so the covariance between $\lg A$ and γ in posterior changes following a rotation of a power law about this frequency. We tested that at $f = 10\text{yr}^{-1}$ our improved model mostly affects the posterior on γ , introducing consistency with 13/3, leaving the posterior on $\lg A_{10\text{yr}}$ almost unchanged. Precisely, with 10-year EPTA data, we measure $(\lg A_{10\text{yr}}, \gamma)$ to be $(-14.02^{+0.08}_{-0.11}, 4.35^{+0.42}_{-1.02})$ with our new noise model and $(-14.00^{+0.08}_{-0.13}, 3.02^{+0.89}_{-0.25})$ with the original model. Consistently, one may notice in Figure A1 in ref. 12 that the posterior on $\lg A_{10\text{yr}}$ referenced to $f = 10\text{yr}^{-1}$ are in less tension with theoretical predictions.

As shown in ref. 16, Equation (1) can be presented as a function of black hole number density ρ_{BH} , black hole mass scale M_* , and the intrinsic scatter ϵ_0 in galaxy stellar velocity dispersion, which is a proxy of mass for supermassive black holes,

$$h_c^2(f) = 1.2 \times 10^{-30} \left(\frac{f}{\text{yr}^{-1}}\right)^{-4/3} \times \left(\frac{M_*}{5.8 \times 10^7 M_\odot}\right)^{2/3} \times \left(\frac{\rho_{\text{BH}}}{4.5 \times 10^5 M_\odot \text{Mpc}^{-3}}\right) \times \left(\frac{2}{e^{\frac{8}{3}\epsilon_0^2 \ln^2(10)}}\right). \quad (2)$$

The mass scale M_* is expressed as $\lg M_* = a_* + b_* \frac{\sigma_*}{200 \text{km s}^{-1}}$. It is also worth noting that $\rho_{\text{BH}} \propto M_*$. Based on the equation above, we perform parameter estimation on $(\rho_{\text{BH}}, a_*, b_*, \sigma_*)$ with the 10-year EPTA data and Hellings-Downs correlations. We assume Normal priors for (a_*, b_*, σ_*) as well as the value of $\epsilon_0 = 0.38$ from ref. 16. The priors for (a_*, b_*) for black hole masses are inferred from kinematic observations of local black holes in galaxy catalogues⁶⁶. The prior for σ_* is based on the velocity dispersion measured in the Sloan Digital Sky Survey⁶⁷. We adopt a Normal prior on ρ_{BH} from ref. 18 because a prior from ref. 16 yields the strain amplitude in significant tension with the EPTA observations. The posterior and the prior are shown in Fig. 3. All parameters are degenerate with each other due to representing a single observed quantity, the strain amplitude. Nevertheless, the figure illustrates which properties of the supermassive black hole population are constrained the most by the pulsar timing data, given astrophysical uncertainties. Provided the scaling in Equation (2), (0.6%, 5.7%, 0.9%) uncertainties on (a_*, b_*, σ_*) , which contribute to M_* , and a larger 44.4% uncertainty on ρ_{BH} , the posterior informs solely on the number density ρ_{BH} . Our improved model influences the posterior to a lesser extent compared to Fig. 1 because the spectral index is fixed at $\gamma = 13/3$. Accordingly, it can be noticed in Fig. 1 that $\lg A$ does not change much for our improved noise model at the slices of fixed $\gamma = 13/3$.

Discussion

When the model closely matches reality, one would expect a reduction of the measurement uncertainty when adding extra data. This is not visible in the original EPTA analysis, where the 1σ range for $(\lg A, \gamma)$ spans 0.44 for the 10-year data and 0.39 for the 25-year data before our

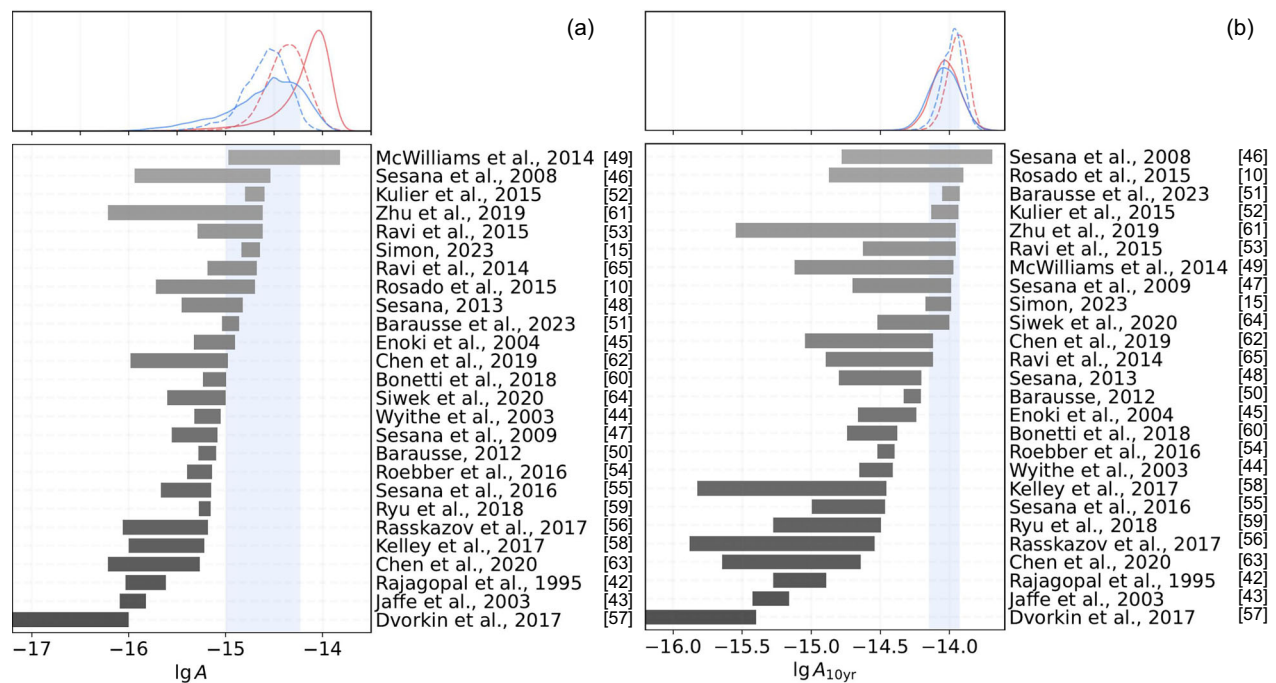


Fig. 2 | Consistency of the inferred strain amplitude of the gravitational wave background with theoretical expectations. Bottom panels show the predicted log-10 strain amplitude $\lg A$ from 26 studies. The top panels show posteriors on $\lg A$ marginalised over γ . Red lines correspond to the original noise model, and blue lines correspond to our revised model. Solid lines correspond to the 10-year data,

and dashed lines correspond to the 25-year data. Blue bands correspond to 1σ credible levels for 10-year data and our improved noise model. Panel (a) shows the characteristic strain amplitude at yr^{-1} , $\lg A$. Panel (b) shows the strain amplitude at the frequency of $(10\text{yr})^{-1}$.

improved noise model. As shown in Fig. 1c, our improved analysis yields 0.95 for the 10-year data and 0.42 for the 25-year data. A larger reduction in the measurement uncertainty when adding ten extra years of data is in agreement with our expectations. A decrease in 1σ uncertainty levels in Fig. 1c compared to Fig. 1b indicates that inter-pulsar correlations in the 10-year data provide additional constraints on the background amplitude and spectral index. A larger maximum-*a posteriori* $\lg A$ in Fig. 1b compared to both Fig. 1a and Fig. 1c may also be driven by inter-pulsar correlations.

The shift of the posteriors in Fig. 1 towards larger spectral indices and smaller amplitudes with our improved analysis suggests that the louder pulsar-intrinsic noise with flatter spectra leaks into our measurement of the background strain spectrum when the four new sources of noise we point out are not modelled. The covariance between $\lg A$ and γ in Fig. 1 is along the line of equal noise power. Furthermore, in the early part of the 25-year data, our improved model may ameliorate instrumental noise and a lack of frequency coverage in addition to better modelling of millisecond pulsar spin noise³³.

Because the 10-year data and the 25-year data are not independent data sets, a high degree of consistency is expected. In the original EPTA analysis, maximum-*a posteriori* $(\lg A, \gamma)$ in the 25-year data differ from those of the 10-year data by (0.29, 0.76). It is visible in red contours across all three panels in Fig. 1. A smaller difference between the best-fit $(\lg A, \gamma)$ in the 25-year data and in the 10-year data of (0.16, 0.76) is achieved with our improved analysis. When not modelling Hellings-Downs correlations (Fig. 1c), the difference between the best-fit $(\lg A, \gamma)$ in the 25-year data and in the 10-year data is only (0.01, 0.06). Because the best-fit $(\lg A, \gamma)$ obtained with only temporal correlations is still expected to match those obtained with temporal and Hellings-Downs correlations, it is also possible that we have not removed all noise model misspecification from the analysis of the EPTA data.

Let us briefly hypothesise about the nature of any other mis-modeled noise. We note that the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) has mitigated a tension of

the background spectral index with 13/3 by adopting the Gaussian process model of the dispersion variation noise⁵, as in the EPTA analysis²³. Therefore, one potential source of a systematic error could be the mismodelling of the pulsar-specific noise that depends on the radio frequency. A very nearby binary is another example^{68–71}. Frequency-wise comparison of the inferred strain spectrum against black hole population synthesis models performed earlier by the EPTA (Fig. 3 in ref. 13) suggests that the deviation from $\gamma = 13/3$ may occur due to excess noise in two frequency bins, $1.3 \times 10^{-8} \text{ Hz}$ and $2.9 \times 10^{-8} \text{ Hz}$. The rest of the spectrum appears to be consistent with $\gamma = 13/3$. The aforementioned potential sources of systematic errors may require better temporal and inter-pulsar correlation models of the data as part of future work.

Finally, we calculate evidence for Hellings-Downs correlations in the EPTA data with the revised noise model. Namely, the Bayes factor \mathcal{B} in favour of the hypothesis that all pulsars contain Hellings-Downs correlated signal with (A, γ) against the hypothesis of the same signal without Hellings-Downs correlations. For 10-year EPTA data, we find $\mathcal{B} = 38$ ($\ln \mathcal{B} = 3.63$). For the original EPTA noise model, we find a higher value of $\mathcal{B} = 59$, consistent with the previously reported value of 60 in Table 5 in ref. 6. Therefore, our improved model has slightly decreased evidence for Hellings-Downs correlations. It is a combination of the increased measurement uncertainty and, potentially, a fraction of the previously-reported evidence due to the propagation of unmodelled noise into evidence. For 25-year EPTA data, we find $\mathcal{B} = 3$ ($\ln \mathcal{B} = 1.17$). For the original EPTA noise model, we find a consistent value ($\mathcal{B} = 3$, $\ln \mathcal{B} = 0.93$), which is similar to the previously reported value $\mathcal{B} = 4$ from Table 5 in ref. 6.

To evaluate the broad form of inter-pulsar correlations in the 10-year EPTA data, we present Fig. 4. It shows the posterior on inter-pulsar correlations, as a function of angular separation between EPTA pulsar pairs, with our revised noise model. The result shows broad consistency with Hellings-Downs correlations. With this, we should point out a caveat that some of the parameter space in correlation

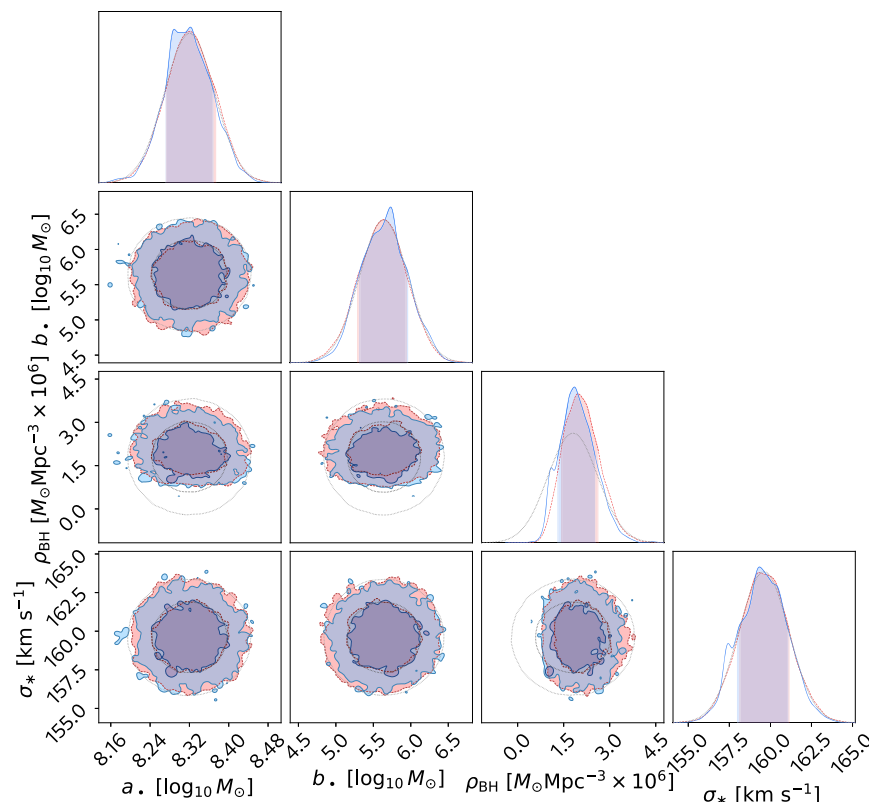


Fig. 3 | Posterior of the population parameters for supermassive black hole binaries with the 10-year data. Effectively, the constraints are provided solely by the inferred characteristic strain amplitude of the gravitational wave background,

assuming the strain power law index of $-2/3$. Hellings-Downs correlations are included in the model.

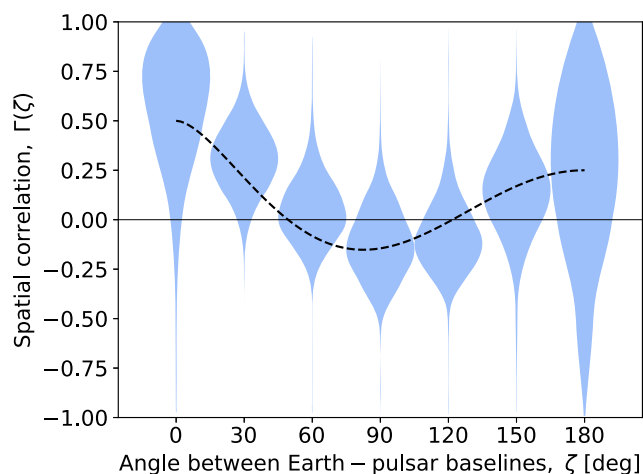


Fig. 4 | Posterior on inter-pulsar correlations of the common stochastic process in the 10-year data. Here, we assume our improved noise model. Correlations are shown as a function of angular separation between the observed pulsar pairs. Hellings-Downs function of the gravitational wave background is shown as the dashed line.

coefficients may be ruled out solely because they do not result in a positive definite covariance matrix of the likelihood. Thus, strictly speaking, flexible correlation models are ill-posed (Rutger van Haasteren, private communication).

Methods

PTAs perform precision measurements of pulse arrival times from millisecond radio pulsars. The likelihood of delays(-advances) $\delta\mathbf{t}$ for a

vector of pulse arrival times \mathbf{t} is modelled as a multivariate Gaussian distribution $\mathcal{L}(\delta\mathbf{t}|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of parameters of models that describe the data. From Bayes' theorem, it follows that the posterior distribution of model parameters is $\mathcal{P}(\boldsymbol{\theta}|\delta\mathbf{t}) = \mathcal{Z}^{-1} \mathcal{L}(\delta\mathbf{t}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$, where \mathcal{Z} is Bayesian evidence, a fully-marginalised likelihood. The term $\pi(\boldsymbol{\theta})$ is called a prior probability distribution, a model of how likely it is to find a certain value of $\boldsymbol{\theta}$ in Nature. Model selection is performed by computing the ratio of \mathcal{Z} for pairs of models, which is referred to as the Bayes factor. The Bayes factor is equal to the Bayesian odds ratio if both models are assumed to have equal prior odds.

The time-domain likelihood of the observed radio pulsar pulse delays-advances time series $\delta\mathbf{t}$, at the position and of the Solar System barycenter, is a multivariate Gaussian,

$$\mathcal{L}(\delta\mathbf{t}|\boldsymbol{\theta}) = \frac{\exp\left(-\frac{1}{2}(\delta\mathbf{t} - \boldsymbol{\mu})^T \mathbf{C}^{-1}(\delta\mathbf{t} - \boldsymbol{\mu})\right)}{\sqrt{\det(2\pi\mathbf{C})}}. \quad (3)$$

Here, $\boldsymbol{\theta}$ are model parameters. Vector $\boldsymbol{\mu}$, a function of $\boldsymbol{\theta}$, represents the model prediction for timing residuals. The covariance matrix is $\mathbf{C} = \mathbf{N} + \mathbf{T}^T \mathbf{B}^{-1} \mathbf{T}$. The likelihood is marginalised over coefficients \mathbf{b} which determine the time series realisation \mathbf{Tb} of the corresponding signal or noise. Square matrix \mathbf{N} only contains diagonal elements that represent temporally uncorrelated noise, which is referred to as white. The so-called design matrix $\mathbf{T} = [\mathbf{M}, \mathbf{F}, \mathbf{U}]$ is made up of three blocks: for the timing model, \mathbf{M} , for the Fourier series of time-correlated signals, \mathbf{F} , and for the epoch-correlated noise, \mathbf{U} . Design matrix maps the model parameter space (columns) to the pulse time of arrival space (rows), such that \mathbf{Tb} is the time series, like $\boldsymbol{\mu}$. Precisely, parameters \mathbf{e} for \mathbf{M} are those of the EPTA timing model⁷², including pulsar spin frequency, its derivative, pulsar sky position, *etc.* Parameters for \mathbf{F} are power spectral

density amplitudes \mathbf{a} corresponding to Fourier sine and cosine terms at Fourier frequencies. We model temporal correlations in 30 frequencies for a common time-correlated stochastic signal attributed to the gravitational wave background; the number of frequencies for time-correlated noise is determined based on single-pulsar noise analysis²³. Parameters for \mathbf{U} are delay time series for every epoch, j ; epochs are groups of nearly simultaneous pulse arrival times at different radio frequencies. Coefficients $\mathbf{b} = [\mathbf{e}, \mathbf{a}, \mathbf{f}]$, corresponding to the same signals as in \mathbf{T} , represent signal amplitudes before a transformation to the time domain. Finally, the covariance matrix of coefficients, \mathbf{B} , is such that the coefficients \mathbf{b} can be generated from the zero-mean Gaussian distribution described by the covariance matrix \mathbf{B} . It is the matrix with diagonal blocks $[\boldsymbol{\xi}, \boldsymbol{\phi}, \mathcal{J}]$. The first component $\boldsymbol{\xi}$, a diagonal matrix of the values of 10^{40} , corresponds to a wide, uninformative prior on the timing model coefficients \mathbf{e} . Component \mathcal{J} is the covariance matrix of white epoch-correlated noise. The second component, $\boldsymbol{\phi}$, is such that

$$\varphi_{(ai),(bj)} = P_{ai}\delta_{ab}\delta_{ij} + \Gamma_{ab}P_i\delta_{ij}, \quad (4)$$

where (a, b) are pulsar indices, (i, j) are frequency indices, P_{ai} is the power spectral density of noise, P_i is the power spectral density of a common time-correlated signal, and Γ_{ab} determines the degree of spatial correlations between pulsars a and b . In the case of an isotropic gravitational wave background from black hole binaries in circular orbits, Γ_{ab} is given by the Hellings-Downs function⁹:

$$\Gamma_{ab}|_{a \neq b} = \frac{1}{2} - \frac{x_{ab}}{4} + \frac{3}{2}x_{ab} \ln x_{ab}, \quad (5)$$

where ζ_{ab} is the sky separation angle for a given pair of pulsars and $x_{ab} = (1 - \cos \zeta_{ab})/2$. In addition, $\Gamma_{aa} = 1$.

Time-correlated noise, which yields stochastic timing delays, can be categorised into achromatic and chromatic. Achromatic noise does not depend on the pulsar's pulse radio frequency. It is also referred to as spin noise because it represents pulsar rotational irregularities. Chromatic noise⁷³ is characterised by an amplitude that depends on the radio frequency at which it is observed. Most EPTA pulsars have measurable levels of dispersion measure (DM) noise, and many pulsars are shown to have spin noise²³. Epoch-correlated noise, associated with pulse jitter (short-term pulse profile variations), can influence the inferred strain spectrum of the gravitational wave background at the highest frequencies because of the "white" nature of this noise. Sudden dips with exponential relaxations in timing delays, associated with pulse shape changes on timescales of days-months³², can also influence parameter estimation for time-correlated signals if not modelled. Noise terms missed in the model, despite evidence, are expected, in many cases, to incorrectly increase the inferred signal amplitude. In particular, the gravitational wave model would attempt to absorb excess noise power.

The priors are our models of how parameters governing pulsar-specific noise are distributed across the pulsars. PTA data provides information about the distribution of $\boldsymbol{\theta}$ in Nature, so failing to model this distribution could lead to prior misspecification. To address this, we parametrise prior distributions for amplitudes and spectral indices of pulsar spin and DM noise spectra. Our improved analysis introduces hyperparameters $\boldsymbol{\Lambda}$ to parametrise priors: $\pi(\boldsymbol{\theta}|\boldsymbol{\Lambda})\pi(\boldsymbol{\Lambda})$. We then perform a numerical marginalisation over $\boldsymbol{\Lambda}$. We use a new procedure of marginalisation over hyperparameters, which is described in Section 2.2.2 of the companion paper²⁴.

Because the gravitational wave background is a stochastic time-correlated signal, it is important to hierarchically model other stochastic time-correlated signals in the data. Pulsar spin noise is described by amplitudes and spectral indices $(A_{\text{SN}}, \gamma_{\text{SN}})$, which are different for every pulsar. Similarly, DM noise is characterised by $(A_{\text{DM}}, \gamma_{\text{DM}})$,

where the amplitude is referenced to 1400 MHz. Hierarchical noise model directly impacts posteriors on $(A_{\text{SN}}, \gamma_{\text{SN}}, A_{\text{DM}}, \gamma_{\text{DM}})$. Because the data is only compatible with a certain total amount of timing fluctuations, hierarchical noise inference also impacts posteriors on the gravitational wave background's amplitude.

Data availability

Posterior samples generated in this study have been deposited in the database at zenodo.org/record/15716517⁷⁵. Second data release of the European Pulsar Timing Array⁷² is available at zenodo.org/record/8091568⁷⁶ and gitlab.in2p3.fr/epta/epta-dr2.

Code availability

The code to marginalise over the uncertainties in pulsar noise priors is available at github.com/bvgoncharov/pta_priors (commit "96cd7cc" is used for the analysis presented in this work). The PTA likelihood is incorporated in ENTERPRISE⁷⁷, and posterior sampling is performed using PTMCMCSAMPLER⁷⁸.

References

1. Nanograv Collaboration. The NANOGrav 12.5 yr data set: search for an isotropic stochastic gravitational-wave background. *ApJL* **905**, L34 (2020).
2. Goncharov, B. et al. On the evidence for a common-spectrum process in the search for the nanohertz gravitational-wave background with the parkes pulsar timing Array. *ApJ* **917**, L19 (2021).
3. Chen, S. et al. Common-red-signal analysis with 24-yr high-precision timing of the European Pulsar Timing Array: inferences in the stochastic gravitational-wave background search. *MNRAS* **508**, 4970–4993 (2021).
4. Goncharov, B. et al. Consistency of the Parkes Pulsar Timing Array Signal with a Nanohertz Gravitational-wave Background. *ApJ* **932**, L22 (2022).
5. Nanograv Collaboration. The NANOGrav 15 yr data set: evidence for a gravitational-wave background. *ApJL* **951**, L8 (2023).
6. EPTA Collaboration & InPTA Collaboration. The second data release from the European Pulsar Timing Array. III. Search for gravitational wave signals. *Astron. Astrophys.* **678**, A50 (2023).
7. Reardon, D. J. et al. Search for an isotropic gravitational-wave background with the Parkes Pulsar Timing Array. *ApJ* **951**, L6 (2023).
8. Xu, H. et al. Searching for the nano-hertz stochastic gravitational wave background with the Chinese Pulsar Timing Array Data Release I. *Res. Astron. Astrophys.* **23**, 075024 (2023).
9. Hellings, R. W. & Downs, G. S. Upper limits on the isotropic gravitational radiation background from pulsar timing analysis. *ApJ* **265**, L39–L42 (1983).
10. Rosado, P. A., Sesana, A. & Gair, J. Expected properties of the first gravitational wave signal detected with pulsar timing arrays. *MNRAS* **451**, 2417–2433 (2015).
11. Mingarelli, C. M. F. et al. Insights into supermassive black hole mergers from the gravitational wave background. *Nat. Astron.* **9**, 183–184 (2025).
12. Nanograv Collaboration. The NANOGrav 15 yr data set: Constraints on supermassive black hole binaries from the gravitational-wave background. *ApJ* **952**, L37 (2023).
13. EPTA Collaboration & InPTA Collaboration. The second data release from the European Pulsar Timing Array. IV. Implications for massive black holes, dark matter, and the early Universe. *Astron. Astrophys.* **685**, A94 (2024).
14. Casey-Clyde, J. A. et al. A Quasar-based supermassive black hole binary population model: Implications for the gravitational wave background. *ApJ* **924**, 93 (2022).
15. Simon, J. Exploring [roxies for the supermassive black hole mass function: Implications for Pulsar Timing Arrays. *ApJ* **949**, L24 (2023).

16. Sato-Polito, G., Zaldarriaga, M. & Quataert, E. Where are the supermassive black holes measured by PTAs? *Phys. Rev. D* **110**, 063020 (2024).
17. Sato-Polito, G. & Zaldarriaga, M. Distribution of the gravitational-wave background from supermassive black holes. *Phys. Rev. D* **111**, 023043 (2025).
18. Liepold, E. R. & Ma, C.-P. Big galaxies and big black holes: The massive ends of the local stellar and black hole mass functions and the implications for nanohertz gravitational waves. *ApJ* **971**, L29 (2024).
19. Renzini, A. I., Goncharov, B., Jenkins, A. C. & Meyers, P. M. Stochastic gravitational-wave backgrounds: current detection efforts and future prospects. *Galaxies* **10**, 34 (2022).
20. Middleton, H., Chen, S., Del Pozzo, W., Sesana, A. & Vecchio, A. No tension between assembly models of super massive black hole binaries and pulsar observations. *Nat. Commun.* **9**, 573 (2018).
21. Afzal, A. et al. The NANOGrav 15 yr data set: Search for signals from new physics. *ApJ* **951**, L11 (2023).
22. Di Marco, V., Zic, A., Shannon, R. M., Thrane, E. & Kulkarni, A. D. Choosing suitable noise models for nanohertz gravitational-wave astrophysics. *ApJ* **990**, 85 (2025).
23. EPTA Collaboration & InPTA Collaboration. The second data release from the European Pulsar Timing Array. II. Customised pulsar noise models for spatially correlated gravitational waves. *Astron. Astrophys.* **678**, A49 (2023).
24. Goncharov, B. & Sardana, S. Ensemble noise properties of the European Pulsar Timing Array. *MNRAS* **537**, 3470–3479 (2025).
25. Zic, A. et al. Evaluating the prevalence of spurious correlations in pulsar timing array data sets. *MNRAS* **516**, 410–420 (2022).
26. Cornish, N. J. & Sampson, L. Towards robust gravitational wave detection with pulsar timing arrays. *Phys. Rev. D* **93**, 104047 (2016).
27. Taylor, S. R. et al. All correlations must die: Assessing the significance of a stochastic gravitational-wave background in pulsar timing arrays. *Phys. Rev. D* **95**, 042002 (2017).
28. Di Marco, V. et al. Toward robust detections of nanohertz gravitational waves. *Astrophys. J.* **956**, 14 (2023).
29. van Haasteren, R. Pulsar timing arrays require hierarchical models. *Astrophys. J.* **773**, 23 (2024).
30. NANOGrav Collaboration. The NANOGrav nine-year data set: Limits on the isotropic stochastic gravitational wave background. *Astrophys. J.* **821**, 13 (2016).
31. van Haasteren, R. Use model averaging instead of model selection in pulsar timing. *MNRAS* **537**, L1–L6 (2025).
32. Goncharov, B. et al. Identifying and mitigating noise sources in precision pulsar timing data sets. *MNRAS* **502**, 478–493 (2021).
33. Ferranti, I. et al. Impact of the observation frequency coverage on the significance of a gravitational wave background detection in pulsar timing array data. *Astron. Astrophys.* **694**, A38 (2025).
34. Lamb, W. G. & Taylor, S. R. Spectral Variance in a Stochastic Gravitational-wave Background from a Binary Population. *Astrophys. J.* **971**, L10 (2024).
35. Allen, B. Variance of the Hellings-Downs correlation. *Phys. Rev. D* **107**, 043018 (2023).
36. Saeedzadeh, V., Mukherjee, S., Babul, A., Tremmel, M. & Quinn, T. R. Shining light on the hosts of the nano-Hertz gravitational wave sources: a theoretical perspective. *MNRAS* **529**, 4295–4310 (2024).
37. Sesana, A. Self consistent model for the evolution of eccentric massive black hole binaries in stellar environments: implications for gravitational wave observations. *Astrophys. J.* **719**, 851–864 (2010).
38. Sesana, A. & Khan, F. M. Scattering experiments meet N-body - I. A practical recipe for the evolution of massive black hole binaries in stellar environments. *MNRAS* **454**, L66–L70 (2015).
39. Bonetti, M., Franchini, A., Galuzzi, B. G. & Sesana, A. Neural networks unveiling the properties of gravitational wave background from supermassive black hole binaries. *Astron. Astrophys.* **687**, A42 (2024).
40. Bi, Y.-C., Wu, Y.-M., Chen, Z.-C. & Huang, Q.-G. Implications for the supermassive black hole binaries from the NANOGrav 15-year data set. *Sci. China Phys. Mech. Astron.* **66**, 120402 (2023).
41. Burke-Spolaor, S. et al. The astrophysics of nanohertz gravitational waves. *Astron. Astrophys. Rev.* **27**, 5 (2019).
42. Rajagopal, M. & Romani, R. W. Ultra-low-frequency gravitational radiation from massive black hole binaries. *Astrophys. J.* **446**, 543 (1995).
43. Jaffe, A. H. & Backer, D. C. Gravitational waves probe the coalescence rate of massive black hole binaries. *Astrophys. J.* **583**, 616–631 (2003).
44. Wyithe, J. S. B. & Loeb, A. Low-frequency gravitational waves from massive black hole binaries: Predictions for LISA and Pulsar Timing Arrays. *Astrophys. J.* **590**, 691–706 (2003).
45. Enoki, M., Inoue, K. T., Nagashima, M. & Sugiyama, N. Gravitational waves from supermassive black hole coalescence in a hierarchical galaxy formation model. *Astrophys. J.* **615**, 19–28 (2004).
46. Sesana, A., Vecchio, A. & Colacino, C. N. The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays. *MNRAS* **390**, 192–209 (2008).
47. Sesana, A., Vecchio, A. & Volonteri, M. Gravitational waves from resolvable massive black hole binary systems and observations with Pulsar Timing Arrays. *MNRAS* **394**, 2255–2265 (2009).
48. Sesana, A. Systematic investigation of the expected gravitational wave signal from supermassive black hole binaries in the pulsar timing band. *MNRAS* **433**, L1–L5 (2013).
49. McWilliams, S. T., Ostriker, J. P. & Pretorius, F. Gravitational Waves and Stalled Satellites from Massive Galaxy Mergers at $z \leq 1$. *Astrophys. J.* **789**, 156 (2014).
50. Barausse, E. The evolution of massive black holes and their spins in their galactic hosts. *MNRAS* **423**, 2533–2557 (2012).
51. Barausse, E. et al. Implications of the pulsar timing array detections for massive black hole mergers in the LISA band. *Phys. Rev. D* **108**, 103034 (2023).
52. Kulier, A., Ostriker, J. P., Natarajan, P., Lackner, C. N. & Cen, R. Understanding black hole mass assembly via accretion and mergers at late times in cosmological simulations. *Astrophys. J.* **799**, 178 (2015).
53. Ravi, V., Wyithe, J. S. B., Shannon, R. M. & Hobbs, G. Prospects for gravitational-wave detection and supermassive black hole astrophysics with pulsar timing arrays. *MNRAS* **447**, 2772–2783 (2015).
54. Roebber, E., Holder, G., Holz, D. E. & Warren, M. Cosmic variance in the nanohertz gravitational wave background. *Astrophys. J.* **819**, 163 (2016).
55. Sesana, A., Shankar, F., Bernardi, M. & Sheth, R. K. Selection bias in dynamically measured supermassive black hole samples: consequences for pulsar timing arrays. *MNRAS* **463**, L6–L11 (2016).
56. Rasskazov, A. & Merritt, D. Evolution of massive black hole binaries in rotating stellar nuclei: Implications for gravitational wave detection. *Phys. Rev. D* **95**, 084032 (2017).
57. Dvorkin, I. & Barausse, E. The nightmare scenario: measuring the stochastic gravitational wave background from stalling massive black hole binaries with pulsar timing arrays. *MNRAS* **470**, 4547–4556 (2017).
58. Kelley, L. Z., Blecha, L., Hernquist, L., Sesana, A. & Taylor, S. R. The gravitational wave background from massive black hole binaries in Illustris: spectral features and time to detection with pulsar timing arrays. *MNRAS* **471**, 4508–4526 (2017).
59. Ryu, T., Perna, R., Haiman, Z., Ostriker, J. P. & Stone, N. C. Interactions between multiple supermassive black holes in galactic nuclei: a solution to the final parsec problem. *MNRAS* **473**, 3410–3433 (2018).

60. Bonetti, M., Sesana, A., Barausse, E. & Haardt, F. Post-Newtonian evolution of massive black hole triplets in galactic nuclei - III. A robust lower limit to the nHz stochastic background of gravitational waves. *MNRAS* **477**, 2599–2612 (2018).
 61. Zhu, X.-J., Cui, W. & Thrane, E. The minimum and maximum gravitational-wave background from supermassive binary black holes. *MNRAS* **482**, 2588–2596 (2019).
 62. Chen, S., Sesana, A. & Conselice, C. J. Constraining astrophysical observables of galaxy and supermassive black hole binary mergers using pulsar timing arrays. *MNRAS* **488**, 401–418 (2019).
 63. Chen, Y., Yu, Q. & Lu, Y. Dynamical evolution of cosmic supermassive binary black holes and their gravitational-wave radiation. *Astrophys. J.* **897**, 86 (2020).
 64. Siwek, M. S., Kelley, L. Z. & Hernquist, L. The effect of differential accretion on the gravitational wave background and the present-day MBH binary population. *MNRAS* **498**, 537–547 (2020).
 65. Ravi, V., Wyithe, J. S. B., Shannon, R. M., Hobbs, G. & Manchester, R. N. Binary supermassive black hole environments diminish the gravitational wave signal in the pulsar timing band. *MNRAS* **442**, 56–68 (2014).
 66. McConnell, N. J. & Ma, C.-P. Revisiting the scaling relations of black hole masses and host galaxy properties. *Astrophys. J.* **764**, 184 (2013).
 67. Bernardi, M. et al. Galaxy luminosities, stellar masses, sizes, velocity dispersions as a function of morphological type. *MNRAS* **404**, 2087–2122 (2010).
 68. EPTA Collaboration. The second data release from the European Pulsar Timing Array V. Search for continuous gravitational wave signals. *Astron. Astrophys.* **690**, A118 (2023).
 69. Valtolina, S., Shaifullah, G., Samajdar, A. & Sesana, A. Testing strengths, limitations, and biases of current pulsar timing arrays' detection analyses on realistic data. *Astron. Astrophys.* **683**, A201 (2024).
 70. Ferranti, I., Shaifullah, G., Chalumeau, A. & Sesana, A. Separating deterministic and stochastic gravitational wave signals in realistic pulsar timing array datasets. *Astron. Astrophys.* **694**, A194 (2025).
 71. Bécsy, B. et al. How to detect an astrophysical nanohertz gravitational wave background. *Astrophys. J.* **959**, 9 (2023).
 72. EPTA Collaboration. The second data release from the European Pulsar Timing Array. I. The dataset and timing analysis. *Astron. Astrophys.* **678**, A48 (2023).
 73. Larsen, B. et al. The NANOGrav 15 yr Data set: Chromatic gaussian process noise models for six pulsars. *Astrophys. J.* **972**, 49 (2024).
 74. Perera, B. B. P. et al. The international pulsar timing array: second data release. *MNRAS* **490**, 4666–4687 (2019).
 75. Goncharov, B. et al. Posterior samples: Reading signatures of supermassive binary black holes in pulsar timing array observations <https://doi.org/10.5281/zenodo.15716517> (2025).
 76. Antoniadis, J. et al. The second data release from the european pulsar timing array i. the dataset and timing analysis <https://doi.org/10.5281/zenodo.8300645> (2023).
 77. Ellis, J. A., Vallisneri, M., Taylor, S. R. & Baker, P. T. Enterprise: Enhanced numerical toolbox enabling a robust pulsar inference suite. Zenodo <https://doi.org/10.5281/zenodo.4059815> (2020).
 78. Ellis, J. & van Haasteren, R. jellis18/ptmcmcsampler: Official release <https://doi.org/10.5281/zenodo.1037579> (2017).
- Gabriela Sato-Polito for clarifying the details of ref. 16. This research was supported in part by a grant NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP). Some of our calculations were carried out using the OzSTAR Australian national facility (high-performance computing) at Swinburne University of Technology. European Pulsar Timing Array is a member of the International Pulsar Timing Array⁷⁴. J.A. acknowledges support from the European Commission (ARGOS-CDS; Grant Agreement number: 101094354). A.C. acknowledges financial support provided under the European Union's Horizon Europe ERC Starting Grant "A Gamma-ray Infrastructure to Advance Gravitational Wave Astrophysics" (GIGA; Grant Agreement: 10116134).

Author contributions

B.G. performed the analysis and paper writing. S.S. found the optimal hierarchical model for time-correlated noise and contributed to the figures. A.S. made suggestions on the astrophysical interpretation of the results. S.M.T. found evidence for epoch-correlated noise in the data, which is included in our revised noise model. Other authors listed alphabetically – J.A., A.C., D.J.C., S.C., E.F.K., K.L., G.S., L.S., and S.V. – have performed pulsar observations or otherwise contributed to the development of the data release used in this work.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-65450-3>.

Correspondence and requests for materials should be addressed to Boris Goncharov.

Peer review information *Nature Communications* thanks Emily R. Liepold, and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025

Acknowledgements

We thank Bruce Allen for contributions to the structuring of the early version of this manuscript. We thank Rutger van Haasteren for discussions about hierarchical Bayesian inference. We thank Emily Liepold and other colleagues for helpful comments on the manuscript. We thank